The fruitful, tangled trees of knowledge

The world's telephone systems are quickly being overgrown by the data networks springing up within and around them

ONE day telecom companies will earn more money carrying data from chip to chip than by carrying voices from telephone to telephone. But by then—quite possibly before the turn of the century—the telecom companies will also have lost their supremacy in communications. Networks owned by companies, universities, governments and housing co-operatives will provide crucial links in shipping information across campuses, cities and continents. The reason is that computers pass information to and fro at rates that would dizzy the most talkative telephone user. Making the most of this loquaciousness requires a new world of networks, ranging in size from a tangle of cables joining a laboratory's workstations to a satellite system linking a multinational's laptops.

The new networks provide magnificent opportunities for innovation. For the most part, they will be public thoroughfares; they are being built, however, not by governments, but by many different companies and individuals, seeking to serve first their own needs, then the needs of others. In this the networks are more like early 19th-century railway systems than they are like roads—though neither analogy does justice to the task that the builders of this new infrastructure face.

The sheer diversity of the networks now being developed makes it hard to see what the finished system will look like. Such foresight is made even harder by the fact that, while data networks and telephone networks are intertwined, data-networkers think of the task of communication in a fundamentally different way from their colleagues in the traditional world of telephones. Both need links between places—wires, fibre-optic cables or radio waves. But the way they use those links differs profoundly and, as yet, irreconcilably.

Pipes and packets

People running telephone networks think of communication in terms of channels and connections. They see their job as creating a channel between two machines, a pipe through which a steady stream of information can flow. To make money, companies charge for a pipe according to its length, the amount of information it can carry (its width) and the time it stays open.

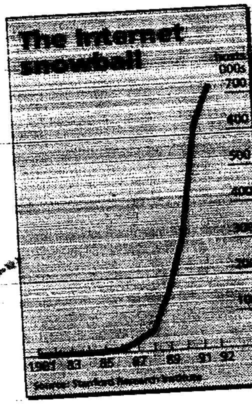
Digital information, though, does not travel in steady streams. It travels in "packets"—more like bits of freight being shipped across country than water flowing through pipes. Think of a message as a large structure—London Bridge, say—that has to be moved from A to B. First it must be broken down into movable pieces. Since all the pieces will be shipped separately—and possibly by different routes—each needs a bill of lading to say what it is and where it should go. The packets sent through a data network also carry such bills, as well as "error-correcting codes" that enable the recipient to see if the data have been damaged in transmission. All this makes things quite complicated—much more so than just letting the data flow down a pipe. But complexity does not matter, as long as there are computers to deal with it and keep it out of sight.

The data networks' hidden complexity gives them great flexibility. The size and shape of a data network are easily changed, more or less at will. Any computer added to the system can help to manage the flow of messages, passing packets on towards their destination by reading their bills of lading. Packets can be sent on their way by whichever route seems best to the computers involved. And they need not arrive in the order they are sent. Networks can be designed so that the receiving computer will take responsibility for reshuffling the packets into order, asking that garbled packets be sent again and generally making up for any deficiencies along the lines.

Systems which can adapt themselves so well to their circumstances do not conform to a single model. The structure of a network is limited only by its technology, its purpose and its history—which is often chequered. Most networks are patchworks, made from available telephone services supplemented by bespoke data-links. Consider a few examples.

Internet is the world's biggest computer network. It links several million people (mostly academics and researchers) around the world through more than 750,000 "hosts". Individual computers are normally linked to a host by small "local-area networks", which use wires owned and operated by universities, companies or whatever. Hosts are linked to each other via long-distance, high-capacity lines leased from telephone companies. Because each host is itself a powerful computer, it can keep con-

stant track of what other hosts it is connected to and how busy the connections seem to be, and then route packets accordingly. Network management is distributed between the hosts.



The lack of any centralised management is so complete that nobody really knows how many hosts there are in Internet. Every three months or so, a computer at Stanford Research Institute asks all the hosts that can hear it to speak up and tell Stanford their names. Every time it receives 20-30% more replies than it did the time before. Recently, many of the new voices have come from hosts overseas, and from the companies that are rapidly growing up around the original heart of Internet—which was paid for by governments and available only to officially sanctioned researchers.

America's **United Parcel Service** is putting together a more controlled, less ambitious patchwork called UPSNET to keep track of the 1m or so packages that it picks up and delivers each day. The network already links 1,300 distribution warehouses in 46 countries. By mid-1993 more than 50,000 delivery lorries in America will be linked to it via mobile telephones. Built with the help of four regional cellular-communications companies, at a total cost of around $500m, UPSNET will be America's first national mobile data network.

For the freer of spirit there is **USENET**, an international network linking millions of people without any central guidance at all. USENET provides electronic mail and a series of "newsgroups" where networkers can debate politics, technology, sex or any other common interest. To join the USENET all you need do is to find a member who will forward messages to your machine—and agree to forward mail and news to others.

May your webs always fuse

Governments on both sides of the Atlantic have looked on these networks and seen that they are good. They want to help. America's government has promised to spend a big chunk of the $1 billion it has earmarked for research on "high-performance" computing on expanding Internet into a faster, more far-reaching "National Research and Education Network". The European Commission is talking about an equally ambitious "European Nervous System", linking governments and research establishments.

Grand ideas, unlikely to do much harm. However, data-networking has grown most vigorously from the bottom up, rather than from the top down. Its stimulus was not high-speed data highways, but the local-area networks that companies and laboratories built so workers could share spreadsheets and researchers pool data. The Yankee Group, a Boston firm that watches the market for information technology, estimates that by 1993 four-fifths of America's white-collar workers will use computers linked to local-area networks. The flow of data over these networks, it believes, is growing by about a third each year, and may soon grow faster. Within a year or so, new chips (from Advanced Micro Devices, among others) may bring down the price of the computer hardware that puts data into packets so much that it will be standard on most personal computers, just as it is already on most high-powered workstations.

Larger networks are born when these lesser webs fuse. Vijay Gurbaxani of the University of California at Irvine has studied the way this process gave birth to BITNET, an international network that links local-area networks in universities so that scientists can share results. He found that young networks tend to grow exponentially. With the addition of each new local-area network, the benefits of joining rise because the network provides more people to talk to. At the same time the costs of joining fall. Newcomers are more likely to find a nearby part of the network to connect themselves to. Improved technology and accumulated experience in forging new links also mean that costs fall as networks grow.

There are, of course, limits to this growth. One is the number of people who might want to join; eventually a network serving a community will saturate it, and grow no more. Another is complexity. A network will, at some point, outgrow the technology that manages it. Absorbing new users and smaller networks will increasingly become a struggle, and the cost of expansion will rise inexorably. BITNET has stopped growing at a relatively modest size. Internet, in contrast, is still snow-balling.

The technology to stop networks becoming unmanageably complicated as they grow is already well advanced. But science and engineering will not be enough by themselves. New commercial and regulatory arrangements will be needed to take advantage of ever bigger networks. These are evolving much more slowly.

The seven-fold way

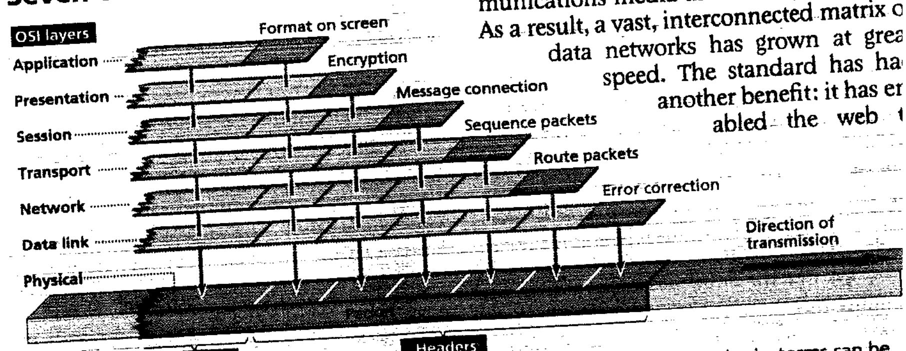
No technology will succeed in the mix-and-match world of network building unless it is governed by accepted standards. At present the design of networks is guided by a sort of meta-standard, the OSI seven-layer model. The OSI model says how data, in their various forms, should look within a network.

At the lowest of the seven levels are standards for sending the 1s and 0s that make up digital data across various media (copper wires, fibre-optic cables, radio waves or whatever). Next are standards that say how those 1s and 0s should look when grouped into packets. At higher levels still are rules about how packets should be assembled into a message. The highest level says how messages should appear on a screen, in the form of paragraphs, pictures and so on. The model is a hierarchy of abstraction: it runs from the nitty-gritty of data on a digital line to the flow of words on a page. What is real to machines is at one end, what is useful to people at the other.

This scheme allows people to disagree about how to do things while agreeing about what to do. A computer will find, stuck on to a travelling packet, tags with the information it needs to deal with the data at each of the seven levels (see diagram). Consider a computer taking a text carried on radio waves and putting it on to a fibre-optic system that works with a completely different sort of packet. The OSI tags let the computer convert the radio-wave dots and dashes into packets, put them in the right order and, going up through the layers, reconstruct the text of the original message. Then, using the second network's methods, it can break the text down into a glimmer of coded light flashing towards some distant machine.

The OSI standard has bridged the gaps between networks based on different com munications media and coding programs As a result, a vast, interconnected matrix o data networks has grown at grea speed. The standard has ha another benefit: it has er abled the web t

Seven views of the same thing



Headers added at each level of the OSI model specify in ever greater detail how brainstorms can be converted to bits—and back again.

keep pace with advances in communications technology.

The faster data move, the more networks can do. Straightforward transmission of a human voice means sending 60,000 or so bits of data a second. An A4 page of text represents about 100,000 bits of data; a photograph about 10m. A cinema-quality film requires a transmission rate of 20m-100m bits a second. That is the sort of rate at which supercomputers can have meaningful conversations, and the true potential of networking can be realised. Happily, fibre-optic cables (though not much else) can move data around at such frightening rates; making use of such spates, though, has its problems.

The walls come tumbling down

In America, supercomputers linked by the fastest parts of Internet can communicate at a theoretical 45m bits a second. Although inefficiencies mean this rate is not quite reached in practice, the network's speed is already comparable to the 20m-80m bits a second at which data typically move about within a single computer. That allows a strange transformation to take place. When computers talk together on a network at the same speed as the component parts of a single computer talk to each other, the computers start to merge. Computers in such a network find, in effect, that they share in the power of all those to which they are linked.

Scientists talk of using such high-powered networks as "virtual laboratories", in which they can work on horrendously complicated problems with far-flung colleagues as though separated by no more than a laboratory bench and a white coat. Visitors to SIGGRAPH, a computer-graphics conference to be held in Chicago in July, will get a feel for this. The National Centre for Supercomputer Applications at the University of Illinois is going to demonstrate BLANCA, its "gigabit testbed".

BLANCA will link supercomputers in central Illinois to workstations on the convention floor over fibre-optic cable transmitting at about 600m bits a second—just over half a gigabit. From the workstations, visitors will be able to peer about within a computer-simulated thunderstorm, created by weather-modelling software. Or they will be able to poke about inside a three-dimensional model of a dog's heart, animated by supercomputer to beat and move just as it does within the dog's body. Or they will be able to tinker with molecular reactions—again using the supercomputer, this time to make the invisible not only visible but also manipulable—using mathematical models based on quantum mechanics.

Later this decade a new technology called "asynchronous transfer mode" (ATM) may bring such capabilities to a broader audience, which will undoubtedly find more practical uses for it. At the moment, high-speed flows of data are hobbled by slow switches at each end. The faster the flow, the less time each computer in the network is granted to decide where a packet should go. ATM networks allow that decision to be made faster by giving the computer less to think about. Unlike today's packets, which come in a variety of different sizes even within each standard, ATM packets are based on a "cell" of standard length. The computer no longer has to think about where the end of each packet might be, which saves valuable nanoseconds.

In addition to the destination of the data, each packet might contain some indication of their urgency. Thus packets representing parts of a videoconference could be hastened along to ensure a smooth picture, while data being dumped in a spreadsheet

file on someone's desk could be held up for a few milliseconds if communications capacity was in short supply.

Lawrence Roberts, one of the creators of Internet, has founded a company called NetExpress which is building big ATM switches in a partnership with DSC Communications, an American telecoms firm. He hopes to begin selling his wares to telephone companies in about 1995. He reckons that telephone companies might by 2000 or so be offering the public data-networking services capable of transmitting at about the same rate as BLANCA. And ATM can also be applied on more modest scales—even down to that of the individual computer. At the University of Cambridge's Computer Laboratory, among other places, researchers are experimenting with ATM to move data within as well as between computers—though that distinction may not mean much in the high-speed-network future.

More than human

With Mr Roberts and his fellow engineers creating the capacity to flood computers with data, software engineers are trying to stop the poor things from being swamped, and from suffering identity crises. An unprepared computer thrown into a network where data are shared, boundaries blurred and resources likely to be called on by strangers at a moment's notice would be as disconcerted as a hermit dumped in a crowd of telepaths. And not only do the software engineers have to keep the computers sane, they also have to make them simple enough for people to use.

This last job is best done by allowing the user to ignore the network he is using, and to deal with it as if it were a single machine. Such deception depends, in part, on the network file system. This enables computers to peruse data from all over the network as if they were sitting on the computer's own disk drive. Also, "remote-procedure-call" technologies allow a program on one computer to co-operate with another on the other side of the network as if they were running on the same machine. Both techniques enable the computers to know what they are doing to whom, and who is doing what to them, thus keeping things straight.

Networks are tools to let users, as well as machines, communicate. A coming attraction is multimedia mail—electronic mail that contains recorded video and sound as well as text. Sun, maker of the market-leading network file system, is offering this on its new workstations. Apple and Microsoft promise something similar for networks of personal computers soon. It is easy to develop because of the OSI standard, which dictates that a packet carrying text be transmitted exactly like one carrying video. Multimedia messages are just like any other, except bigger.

Networks also allow vast amounts of data to be searched. Brewster Kahne at Thinking Machines, a computer company in Cambridge, Massachusetts, is trying to build a "Wide-Area Information Service". By creating standards for describing the information a particular computer offers, and for queries about that information, it should allow a student to search databases around the world as easily as he could search a computer at his local library. Researchers at CERN in Geneva are trying to build a similar service, which they call World-Wide Web (or w3), drawing on their experience in helping physicists find their way around the mountains of data produced by CERN's accelerators.

Meanwhile other researchers are working on systems that allow one network to find out what another offers. Then they could know the answers to questions that users might ask. Examples: "Does that network have machines that can display video?"; or "Is Marcel Dupont on this network?" One popular approach, led by a consortium of computer companies called the Object Management Group, is to use a fashionable technique called object-oriented programming. The idea is to make everything—every piece of hardware, every program, every definable sub-routine, every cache of data, everything—on the network capable of sending, or responding to, a cer-

tain set of messages. The hard part is to decide what this universally understandable set of messages ought to consist of.

A complementary approach is being developed by a small British company called Architecture Projects Management (under contract to many of the same companies that make up the OMG). It is to build into each network a "trader" that can advertise its capabilities to other networks, and manage its own use of the other networks' similarly advertised services. With such capabilities, networks can form little automated economies, buying and selling disk space, computing power and other services from each other—within limits set by their human masters. Experiments, most notably those at Xerox PARC and Stanford, suggest that such markets may be the best way to manage the explosion of network capabilities.

Trust me

The vision emerging from the research laboratories is one of computers sharing work and data, unhindered by the constraints of space. By co-operating intimately they will provide services none could offer on its own. Back in the real world, however, two practical considerations mar this vision: money and security. People will want to be sure that their machine is sharing its secrets with friends rather than enemies. Those providing data networks will want to be paid for their services—as will, most likely, those whose computers use the networks and thereby end up providing services for others.

Security is an especially frustrating obstacle in the path to network growth. Techniques for so-called public-key cryptography offer relatively easy-to-use ways of creating nearly unbreakable codes to protect data, and nearly unforgeable "digital signatures" to show that a document purporting to come from John Hancock really does. One of the better public-key cryptography techniques has been patented by a small Silicon Valley firm called RSA Data Security and licensed by a host of computer companies such as Sun, Apple, Microsoft and DEC. But America's government, which could do much to establish the necessary standards for data-security technology, seems unwilling to endorse this technology—or any other.

The federal government has repeatedly delayed the choice of a public-key cryptography system for its own use (which would create a *de facto* national standard). It also bans the export of most data-security software and hardware—which means that international networks cannot use systems like RSA. This may stop foreign governments using American technology to put their communications beyond the reach of American spies, but it also prevents multinational companies from using encryption to improve their data networks. Meanwhile, at home, agents from the Federal Bureau of Investigation have been skulking about Congress trying to insert a "right-to-tap" amendment into various communications bills, which would ban criminals from unbreakably encoding evidence of their malfeasance. Where does that leave the fifth amendment?

Confusion over technology, however, pales beside the greater confusion over how to make a business of networking. Most data-workers would agree that their present arrangements with telephone companies are unsatisfactory. The leased lines that provide the long-distance backbones of their services are inflexible and expensive—particularly in Europe, where lack of competition means that leased lines cost up to ten times more than they do America. Worse, because data traffic is so "bursty", data-networkers have to support leased lines that work, on average, at only 1-3% of capacity.

Data-networkers would like to buy more suitable services from the telephone companies. Telephone companies, tied to the traditions of voice service, have failed, by and large, to offer those services. There are some encouraging signs of innovation, especially in America. US West, a regional telecom company, is conducting an experiment in which firms create local-area networks by plugging computers into wall-sockets owned and operated by US West rather than buying their own equipment. Other companies are offering new services to carry data in bulk more cheaply and flexibly; one such is ISDN (integrated services digital networks), which provides digital lines capable of transmitting up to 2m bits a second. It is still expensive and often unavailable, though, and its speed, now impressive, could seem a bit sluggish by the time it becomes widespread.

It is on the basis of such services that American telephone companies think they can increase the money their data-networking services earn, from $4 billion a year today (10% of their total business) to as much as $40 billion a year (50% of the total business) in ten years' time. That growth will precipitate wide-ranging regulatory change. In America and elsewhere, telephones and cable television are kept well apart. Videoconferencing, multimedia mail and the like fit into the gap between them. Are telephone companies really to be allowed to carry television for IBM and GM, but not for NBC and CBS?

More puzzling than who does it, though, is how to pay for it. Today's pricing schemes, which concentrate on connections and time, are ill-fitted to data networks. Data move about all the time, but some need to be kept flowing more than others. Helmut and Fatima's video-conference packets are more time-sensitive than the e-mail going from Carol to Nguyen. Should pricing be based on the number of packets, or on the number of packets plus some measure of speed and quality of service, or on something else?

Such commercial details may well determine the future of the world's communications system. But they are unlikely to cramp its growth unduly. One reason is that the technology already exists to create wonders. These wonders are in demand, so someone will find a way to make the technology pay. Another is that networks are, first and foremost, flexible. They can adapt themselves to any environment, even the worst that uncooperative companies and tardy regulators can provide. It may take time, but one day the computers of the world will unite.